Physica A 565 (2021) 125572

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Reciprocal spreading and debunking processes of online misinformation: A new rumor spreading–debunking model with a case study☆

Meiling Jiang [a,b,1], Qingwu Gao [a,c,1], Jun Zhuang [a,*]

[a] Department of Industrial and Systems Engineering, University at Buffalo, Buffalo, NY, USA
[b] Department of Management Science and Systems, University at Buffalo, Buffalo, NY, USA
[c] School of Statistics and Mathematics, Nanjing Audit University, Nanjing, China

A R T I C L E I N F O

Article history:
Received 9 May 2019
Received in revised form 23 October 2020
Available online 1 December 2020

Keywords:
RSD model
Online misinformation
Interplay of spreading and debunking
Decision making
Hurricane Harvey

A B S T R A C T

In this digital era, massive digital misinformation was ranked first by the World Economic Forum among the top future global risks. As human and financial resources are limited, governments or companies would like to use the optimal level of debunking effort and the most efficient debunking strategy. There exists a rich literature that studies the rumor spreading process on social media. However, a huge gap exists on studying the simultaneous propagation of false rumors and debunking information, and the interplay between them. The spreading of rumors and anti-rumors is a dynamic and reciprocal process. Acknowledging that effective debunking strategy is a potential tool to reduce the loss of massive digital misinformation, this paper proposes a novel rumor spreading–debunking (RSD) model by ordinary differential equation (ODE) system to explore the interplay mechanism between rumor spreading and debunking processes. We derive and discuss the key factors and parameters that influence the debunking process. Firstly, we consider the spreading pattern of a rumor before Debunkers appear based on the Susceptible–Infected–Recovered (SIR) model with its own characteristics of rumor, and obtain a series of results including the final scope of the rumor spreading, the maximal scale of the rumor spreader, the number of Stiflers at any time point, and the popularity level of the rumor. Secondly, with the data from the real world rumor case, which is the "Immigration Rumor" during Hurricane Harvey in 2017, we determine the case-specific parameters, and validate our model by comparing the simulated curve with the real data. Our model helps to understand the impact of the rumor on the social media, and predict the future trend. Finally, we use our model to simulate the influence of different debunking strategy, and identify more efficient debunking measures that should be used by the government officials or companies when facing rumor mill under different situations.

☆ This research was supported by the National Science Foundation (NSF), USA under Award No. 1762807. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. We also thank the editors and two anonymous referees for providing constructive comments.

* Correspondence to: Department of Industrial and Systems Engineering, The State University of New York at Buffalo, 317 Bell Hall, Buffalo, NY 14260-2050, USA.

E-mail address: jzhuang@buffalo.edu (J. Zhuang).

[1] Meiling Jiang and Qingwu Gao are co-first authors.

https://doi.org/10.1016/j.physa.2020.125572

1. Introduction

In today's world, online social media plays a vital role in the information diffusion. With the arrival of the social media, the patterns, role and impact of word-of-mouth have evolved, and new forms of communities were shaped. In the social media age, word-of-mouth can travel faster [1], and opinions may become more asymptotically clustered [2]. The scale, speed and real-time nature of the social Web and the increased possibilities for people to share and express themselves play a crucial role in this [3]. Information diffusion on social media is a double-sided sword. It can be used by authorities for effective information distribution and events management, or by malicious entities to spread rumors and fake news.

Propaganda and misinformation have been used throughout history to influence public opinion. In this digital era, massive digital misinformation was ranked first by the World Economic Forum among the top future global risks, along with water supply crisis, major systemic financial failures, and failure to adapt to climate change [4]. Social media sites such as Facebook and Twitter are reported to be major platforms used to spread fake news [5]. The content of the misinformation includes various topics, such as politics, disasters, economics, entertainment, and sports. For example, on April 23, 2013, a false rumor about a terror attack on the White House, in which President Obama was allegedly injured, provoked an immediate crash in the stock market. Another example is that, during Hurricane Harvey in 2017, false rumors, such as "mandatory evacuations are underway in the City of Houston" [6], "Immigration status has to be checked before you are allowed to enter a shelter" [6], and "residents could not return to the coastal city until all critical services were restored" [7], caused great confusion, panic, and anger among people in the affected area. Great loss of life and property might be caused if people without legal immigration status dare not to enter a shelter for self-protection. During this rumor propagation event, a large amount of Twitter users play an important role both spreading and debunking the misinformation [8].

When a rumor starts circulating on the social media, the first thing the stakeholders would like to do is to debunk the rumor and reduce the loss as soon as possible. Human and financial resources are required for the information management. Typically, once the rumor was brought to the forefront, anti-rumors will be posted on official media immediately, such as official websites, Twitter accounts, and TV news. However, if the scope of the rumor propagation is too wide, the influence of the static anti-rumor on one official media may not be enough. Multiple further debunking strategies could be applied through the social media, for example, by increasing the number of accounts to disseminate the anti-rumors, attracting more people to re-tweet the anti-rumor, or encouraging existing rumor spreaders to delete or clarify the rumors.

As human and financial resources are limited, governments or companies would want to use the more efficient debunking strategy. The spreading of rumors and anti-rumors is a dynamic and reciprocal process. The debunking strategies could be affected by some key factors. The first is the current spreading scope of the rumor. If the rumor is just beginning to circulate in a small scope, less debunking effort is needed. The second is the popularity of the rumor. For some deceptive rumor that are like real and with astonishing content, the spreading could be surprisingly fast. Under such circumstance, more debunking effort should be made. The third is the effectiveness of different debunking strategies. As the interplay between rumor spreading and debunking is dynamic, the effectiveness of the strategies could be different for rumor with different scope and popularity. Therefore, it is imperative to understand the underlying dynamics of the interplay between rumor and anti-rumor, and to investigate the efficiency of different debunking strategy.

Acknowledging that the effective debunking strategy is a potential tool to reduce the loss of massive digital misinformation, this paper proposes a novel model to explore the interplay mechanism between rumor spreading and debunking processes. Specifically, this study attempts to answer three research questions when the government or an organization are in the trouble of rumor mill:

- **RQ1:** What is the current scope, popularity level, and controllable performance of the rumor?
- **RQ2:** What is the more efficient debunking strategy for the rumor at current stage?
- **RQ3:** What is the predicted scope and duration of the rumor, and how do these values vary under different debunking strategies?

Specifically, we use a mathematical method to model the dynamic spreading–debunking process. Rumor spreading is similar to infectious disease spreading, one person could not get to know the rumor unless he or she saw the rumor posted by others. The online rumor spreading would fit into the epidemic modeling even better than the oral spreading, because the underlying social network is relatively tractable [9]. Hence, most of the existing rumor spreading models are built based on epidemic models [10]. In this paper, ordinary differential equations (ODEs) are applied to describe the dynamic spreading–debunking system [11]. The key characteristics of the social network, rumor, and anti-rumor are parameterized in the system of ODEs. With the data from the real world rumor cases, which is the 'Immigration Rumor' during Hurricane Harvey in 2017, we design a Rumor spreading–debunking (RSD) work-flow to determine the case-specific parameters. Then we use our model to simulate the influence of different debunking strategy, and identify the effective debunking efforts under various circumstances.

The rest of paper is organized as follows. In Section 2, we review the existing literature on rumor spreading and online information dissemination. In Section 3, we present our model formulation and analytic propositions. Numerical results and case analyses are presented in Section 4, followed by discussion and concluding remarks in Section 5. The related proofs of this study are available in Appendix.

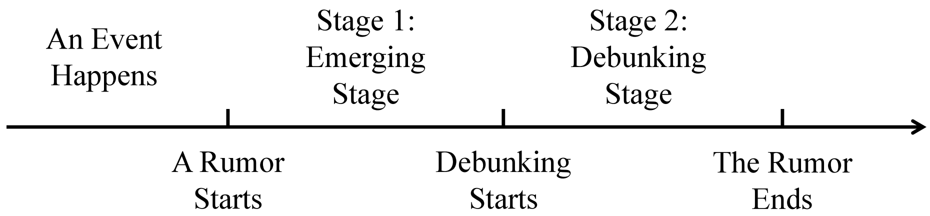


Fig. 1. The life-cycle of the Spreading and Debunking process of a rumor.

2. Literature review

Rumor is a tall tale of explanations of events circulating from person to person and pertaining to an object, event, or issue in public concern [12]. Existing literature have widely varying definitions of rumor. Knapp [13] defined rumor as "a proposition for belief of topical reference disseminated without official verification". The key word in this definition is "unverified". Similarly, Harsin [14] identified rumor as "a claim whose truthfulness is in doubt and which often has no clear source even if its ideological or partisan origins and intents are clear".

The factors that influences the spreading of rumor are various, such as spreading media [15], spreading network topology [16], information characteristics [17], event uncertainty [18], topic type [19], and the interplay among these factors [20]. Moreover, existing literature investigated the rumor spreading process from different angles, such as rumor detection [21], prevention [22], dissemination [23], debunking and intervention [24], rumor source tracking [25]. Some specific users may play an influential role in the information spreading [26]. The spreading of rumor has been studied in multiple disciplines and analyzed from multiple aspects. In this paper, we focus on the interplay between rumor spreading and debunking behaviors, which is still a gap in existing literature. Specifically, we study how the characteristics of the social network, rumor content, and various debunking strategies impact this interplay process.

The dissemination of rumors is a person-to-person informational contagion process [27]. It is natural to mathematically model rumor transmission based on the theory of epidemics [28]. The earliest references in dynamic rumor spreading are based on the Susceptible–Infected–Recovered (SIR) epidemic model, in which the population is stratified into three health states: susceptible state (denoted by S) represents the ones that are susceptible to the infection of the pathogen, infected state (denoted by I) includes those that are infected by the pathogen, and recovered state (denoted by R) refers to those that have been recovered from the infection [29]. Two classical SIR-based models that describe the spread of a rumor were introduced by Kendall et al. [30] and Maki et al. [31] for closed, finite, and homogeneously mixing populations. In the SIR-based rumor propagation model, the population is also divided into three classes: Ignorants (similar to Susceptible, people who never heard the rumor), Spreaders (similar to Infected, people who spread the rumor), and Stiflers (similar to Recovered, people who know the rumor but are not spreading it to others), where the rumor is propagated according to the transition rate among these three classes [32].

There has been one stream of research on the extension of the SIR-based rumor spreading models. Zanette [33] established a rumor spreading model based on small-world networks and provided a threshold of rumor spreading. Zhao et al. [10] extends the classical SIR rumor spreading model by adding a forgetting mechanism. Wang et al. [34] investigated a case when two or more kinds of rumors spread at the same time. However, to our best knowledge, none of the existing literature considered the simultaneous propagation of false rumors and debunking information, and the interplay between them. An initial false report can be circulated very widely if lacking efficient debunking strategies. Therefore, we are motivated to propose a novel interplay model based on the classical SIR rumor propagation model, exploring the interplay mechanism between rumor spreading and debunking processes, and proposing the more efficient debunking strategy.

3. Spreading–debunking competitive model

The life-cycle of a rumor could be divided into two stages. The first stage is the emerging stage, which refers to the time period between the rumor starting point and official debunking point. Various user behaviors exists in this stage, including spreading, questioning, seeking confirmation in the social media [35]. Yet, for most of those who see the rumor, they simply read the information without any posting behavior.

The second stage refers to the time period between the point when the first official anti-rumor appears and the point when the rumor ends. In the second stage, because of the release of official debunking information, people start to spread the anti-rumor in the social network. The number of rumor spreaders starts to decrease. We define the ends of the life-cycle as the point of time when the number of rumor spreaders become under 1. The life-cycle of a rumor is shown in Fig. 1.

In stage 1, the population in the social network consists of three types of people, namely Ignorants, Spreaders, and Stiflers. In stage 2, a new type of people, Debunkers, appear. Debunkers are those who spread the anti-rumor recently. If someone reposts both rumor and anti-rumor recently, we define him as a Debunker. Moreover, in stage 2, there are two type of Stiflers: those who know about the rumor but did not spread the rumor recently (referred to as Stifler_r), and

Table 1
Definitions of the five states.

State	Definition
Ignorant	People who never heard about rumor
Spreader	People who spread the rumor and are contagious
Debunker	People who spread the anti-rumor and are contagious
Stifler_r	People who know about the rumor but do not have contagiousness
Stifler_d	People who know about the anti-rumor but do not have contagiosity

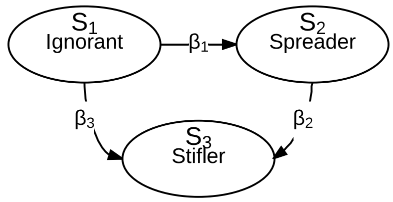


Fig. 2. The transition among Ignorant, Spreader, and Stifler on Stage 1.

those who know about the anti-rumor but did not spread the anti-rumor recently (referred to as Stifler_d). If someone knows about both rumor and anti-rumor but did not spread any recently, we define him as a Stifler_d. Following Kendall et al. [30], we assume the homogeneity of the population. In this way, the rumor spreading and debunking process could be modeled in a macro level. We summarize the definitions of the states of people in Table 1.

In the social media, the role of one individual change over one rumor life cycle. We use an example to illustrate the process. In the very beginning, everyone is the Ignorant, except the initial spreaders (who is the rumor creator). When one individual, say John, saw the rumor posted by his friend on his time-line, he chose to spread it. At the time he reposted the rumor, he became a Spreader. John's friends could see his post and further spread the rumor. Due to information overload, as time goes by (e.g., after one day), people would stop reading and reposting John's post. Hence, we assume that, after one day, he was not a Spreader any more, because he was not contagious to others. At this point of time, he became a Stifler_r. Later, he saw his another friend posted an anti-rumor information to debunk the previous rumor. He reposted the anti-rumor and became a Debunker. Finally, after one day, he lost contagiousness and became a Stifler_d. During the life cycle of this specific rumor, the role of John changes as follows: Ignorant → Spreader → Stifler_r → Debunker → Stifler_d. In this example, only under two states, i.e. Spreader and Debunker states, do people have the ability to infect others. For John, the duration of the Spreader and Debunker states are the one day after which he reposted the rumor or the anti-rumor. We then study the rumor propagation behavior in the two stages, respectively.

3.1. Stage 1: RSD modeling of emerging stage (before Debunkers appear)

In this section, we study the spreading pattern of a rumor before Debunkers appear. We define the initial debunker as the officially verified account who firstly posted the anti-rumor on social media. Before the appearance of the first debunker and/or verification information, the system simply consists of Ignorants, Spreaders, and Stiflers. The spreading scope of rumor at this stage depends on the transition rate among the three states.

We use the transition diagram to describe the transformation among different states of people, see Fig. 2. In this period, the entire population is divided into three groups: Ignorants (S_1), Spreaders (S_2), and Stiflers (S_3). The values of S_1, S_2, and S_3 are all between 0 and 1, indicating the fractions of the population that are in these states. At the initial time point ($t = 0$), the whole population are all Ignorants, except the initial spreaders. The initial spreaders are individuals who firstly spread the rumor intentionally or unintentionally. When Ignorants see the post of the initial Spreaders, some of them repost the rumor and become Spreaders, the others choose to do nothing and become Stiflers. In the transition diagram, β_1 is the spreading rate. When one Ignorant contacts a Spreader, the Ignorant becomes a Spreader with the rate β_1. The spreading rate depends on multiple factors, such as the importance, anxiety, popularity of the rumor content [18]. As all Spreaders will gradually loss contagiosity over time, one Spreader becomes a Stifler with the rate β_2. We name this rate β_2 as the fading rate. When an Ignorant contacts a Spreader, the Ignorant becomes a Stifler with a rumor impression rate β_3. The underlying reason might be that the Ignorant do not believe in the rumor, not be interested in it, or simply do not want to reposted it. The percentages of the population of Ignorants, Spreaders, and Stiflers at time t are $S_1(t)$, $S_2(t)$, and $S_3(t)$, respectively. They satisfy the normalization condition: $S_1(t) + S_2(t) + S_3(t) = 1$.

The rules of rumor propagation are as follows. When a Spreader contacts an Ignorant, the Ignorant may become a Spreader with the rate β_1. Thus, in the equation system, the percentage of people who transit from Ignorant state to Spreader state per time unit is $\beta_1 k S_1(t) S_2(t)$. When a Spreader contacts an Ignorant, the Ignorant may choose not to repost the rumor, and become a Stifler with the rate β_3. The percentage of people who transit from Ignorant state to Stifler state per time unit is $\beta_3 k S_1(t) S_2(t)$. Overtime, the spreader loses contagiousness because of information overload. The percentage of people who transfer from Spreader state to Stifler state per time unit is $\beta_2 S_2(t)$. The index k is the average degree of the network, which is closely related to the density of the network. We use the non-linear ordinary differential equation system to describe the process as follows:

$$\begin{cases} \dfrac{dS_1(t)}{dt} = -\beta_1 k S_1(t) S_2(t) - \beta_3 k S_1(t) S_2(t), \\ \dfrac{dS_2(t)}{dt} = \beta_1 k S_1(t) S_2(t) - \beta_2 S_2(t), \\ \dfrac{dS_3(t)}{dt} = \beta_3 k S_1(t) S_2(t) + \beta_2 S_2(t). \end{cases} \tag{1}$$

In this stage, we are interested in studying the following issues:

(1) Predicting the final scope of the rumor spreading. In other words, without being intervened by the Debunkers, how many people in the social network will be finally impacted by the rumor? Moreover, we study the key factors that influence the final scope.

(2) Predicting the maximal scale of the rumor spreader. In other words, our model tried to predict when would the rumor propagation process arrives the peak, and what will be the maximal number of spreaders that exist in the social network?

(3) Calculating the number of Stiflers at any time point. When a rumor is detected, our model helps officials to calculate how many people have been already aware of the rumor (not necessary spread the rumor) in the social network.

(4) Estimating the popularity level of the rumor. When a rumor is detected, our model helps officials to estimate how fast the rumor is spreading in the social network.

Without debunking effort, some rumors will disappear naturally with very little influence, while some others spread fast and become known by everyone. Hence, we are interested to investigate the expected final scope of the rumor. In other words, without being intervened by Debunkers, how many people will know about the rumor at the end? Let R be the final scale of the rumor, i.e., the fraction of Stiflers at the end, $R = S_3(\infty) \in (0, 1)$. By investigating the equation system (1), we draw the two conclusions below. Because of the technical complexity of this problem, formal proofs are deferred to Appendix.

Proposition 1. *Without debunking intervention, R satisfies the following equation*

$$R = 1 - e^{-\frac{\beta_1 k}{\beta_2} R}. \tag{2}$$

If $\frac{\beta_1 k}{\beta_2} > 1$, *Eq. (2) has two solutions: zero and non-trivial solution R,* $0 < R < 1$.

Proposition 2. *Let all variables of non-trivial solution R be constant except only one, we have that R increases in* β_1 *and k, and decreases in* β_2.

Remark 1. The above two propositions of R provides us some insights of the rumor spreading on the first stage. Proposition 1 derived the implicit equation of the final scope R. This result indicates that the final size of the rumor R is only dependent on three factors: the rumor spreading rate (β_1), fading rate (β_2), and the network degree (k). The value of solution R and its relation with β_1, β_2, and k is investigated and shown in Fig. 3(a). In summary, when $\frac{\beta_1 k}{\beta_2} > 1$, the final scale of the rumor can be non-zero, which means the rumor would circulate among the population if the spreading rate is greater than the fading rate; when $\frac{\beta_1 k}{\beta_2} < 1$, the final scale of the rumor is 0, which means the rumor would not start circulating if the spreading rate is less than the fading rate. The 3D plot and the contour plot in Fig. 3(b) and (c) further illustrated how the final scale R vary over β_1 and β_2.

It is shown in Fig. 3(a) that the parameter $\frac{\beta_1 k}{\beta_2}$ plays a key role to determine the final scope of the rumor. We summarize this feature as follows:

Proposition 3. *Without debunking intervention, the rumor spreading dynamics will remain under control when* $\frac{\beta_1 k}{\beta_2} < 1$ *and uncontrollable when* $\frac{\beta_1 k}{\beta_2} > 1$.

Remark 2. The parameter $\frac{\beta_1 k}{\beta_2}$ is called the rumor spreading control coefficient. The rumor spreading control coefficient is an important and useful parameter in rumor spreading model analysis in sense that it provides information about the rumor spreading process, and it can be used to predict whether the rumor will propagates among the population or not. This coefficient corresponds to the threshold, known as the basic reproduction number in epidemic disease model analysis, see [29]. Proposition 3 suggests when $\frac{\beta_1 k}{\beta_2} > 1$, the rumor invades the population; when $\frac{\beta_1 k}{\beta_2} < 1$, the rumor will gradually disappear, and then debunking interventions may not be needed. The similarity between the rumor spreading mechanism and infectious disease transmission dynamics were highlighted.

Besides the final scale of the rumor, we are also interested in studying the maximum scale of the rumor spreaders. In other words, during the rumor propagation process, how many people are spreading the rumor? We define $S_{\max}$ as the maximum scale of the spreaders, which satisfies the following proposition.

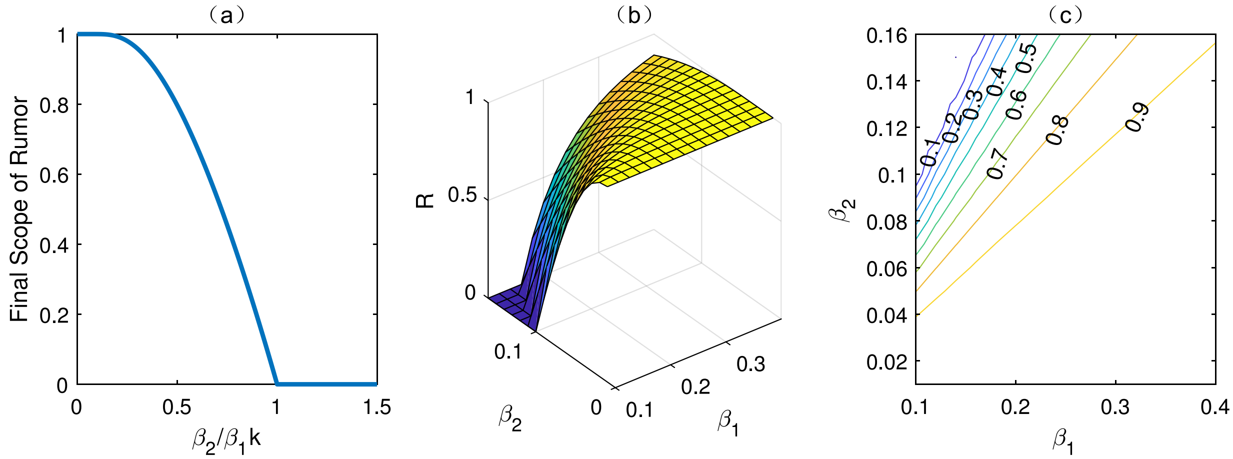


Fig. 3. Analysis of final scale of rumor R.

Proposition 4. *Without debunking intervention, when $S_1 = \frac{\beta_2}{\beta_1 k}$, the system reaches the maximum scale of the spreaders, which is*

$$S_{\max} = \frac{\beta_2}{(\beta_1+\beta_3)k}\left(\ln\frac{\beta_2}{\beta_1 k} - 1\right) + \frac{\beta_1}{\beta_1+\beta_3}.$$

Clearly, we obtain from Proposition 4 that at the point when the Spreaders reach their maximum scale, the percentage of Stiflers is $S_3 = 1 - \frac{\beta_2}{\beta_1 k} - S_{\max}$. The third and fourth research questions at the stage 1 is analyzing the current scope and popularity level of the rumor. In other words, how many people are aware of the rumor, and how fast the rumor is spreading? In the real world, the government can only know how many people have re-posted the rumor from the social media data. However, the majority of the population who know about and are impacted by the rumor may not have any observable spreading behavior. It is unclear how many people have seen the rumor on the network. Moreover, it is hard to directly observe how fast the rumor will keep spreading. In Section 4.1, we demonstrate our solution to this problem by numerically solving our RSD system using real world case study. We design an implementation work flow to derive the optimal solution. The case study validates our model with real life data. Finally, our model could predict the future trend of the rumor spreading.

3.2. Stage 2: RSD modeling with Debunkers

In this section, we discuss the second stage of a rumor event. In this stage, we seek to study the interplay of the rumor and the anti-rumor. Specifically, we are interested to investigate how the final scope and duration of the rumor be impacted by various factors. The factors include the anti-rumor popularity, number of initial Spreaders/Debunkers, intervention time, and so on.

Different from the stage 1, the spreading of the rumor at this stage is impacted by the spreading of the anti-rumor, i.e., the debunking information. In stage 2, the rumor Spreaders and Debunkers are both "contagious" to the Ignorant group. A Debunker can also infect a Spreader, i.e., transit a Spreader to a Debunker by showing them the anti-rumor. Another group of people is called Stifler, who did not spread anything but know about the rumor. If someone know about the rumor, they either only know the rumor, or are already debunked. Thus, in stage 2, we divide the Stiflers into two subgroups: Stiflers_r and Stiflers_d. Specifically, the Stiflers_r refer to people who know about the rumor but did not spread it, or spread it a long time ago and have already lost their influence. Similarly, Stiflers_d mean people who know about the anti-rumor but do not have influence on the others at the focal time point.

See Fig. 4, which indicates the nine sub-processes of the second stage of our RSD model. Consider a closed and homogeneous population consisting of N individuals in a social network. In this stage, the entire population is divided into five groups: Ignorants (S_1), Spreaders (S_2), Debunkers (S_3), Stiflers_r (S_4), and Stiflers_d (S_5). The Contact Eqs. (3)–(11) illustrate this multi-infective propagation process of the RSD model as follows:

$$S_1 + S_2 \xrightarrow{\beta_1} 2S_2, \tag{3}$$

$$S_2 \xrightarrow{\beta_2} S_4, \tag{4}$$

$$S_1 + S_2 \xrightarrow{\beta_3} S_4 + S_2, \tag{5}$$

$$S_1 + S_3 \xrightarrow{\beta_4} S_5 + S_3, \tag{6}$$

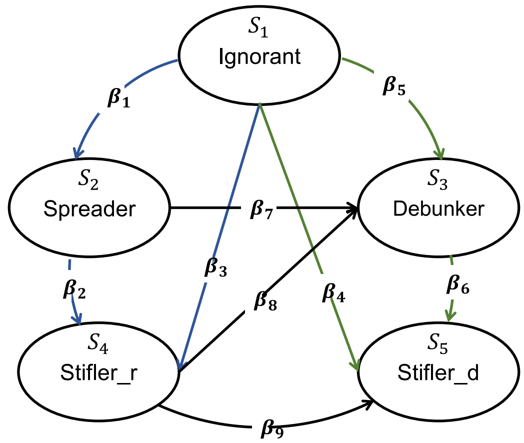


Fig. 4. Transition diagram of Stage 2.

$$S_1 + S_3 \xrightarrow{\beta_5} 2S_3, \quad (7)$$

$$S_3 \xrightarrow{\beta_6} S_5, \quad (8)$$

$$S_2 + S_3 \xrightarrow{\beta_7} 2S_3, \quad (9)$$

$$S_4 + S_3 \xrightarrow{\beta_8} 2S_3, \quad (10)$$

$$S_4 + S_3 \xrightarrow{\beta_9} S_5 + S_3. \quad (11)$$

These Contact Equations describe all possible dynamics that could happen during stage 2. Each equation can be explained as follows:

- Eq. (3): When an Ignorant see the rumor, the Ignorant becomes a Spreader with a rumor spreading rate β_1.
- Eq. (4): As time goes by, the rumor Spreaders lose their influence. The rumor that they spread is no longer noticed by others because of the information overload on social media. Thus, a Spreader becomes a Stifler_r with a rumor fading rate β_2.
- Eq. (5): When an Ignorant see the rumor, the Ignorant may choose not to spread it and become a Stifler_r with a rumor impression rate β_3.
- Eq. (6): When an Ignorant see the anti-rumor, the Ignorant becomes a Stifler_d with a debunking impression rate β_4.
- Eq. (7): When an Ignorant see the anti-rumor, the Ignorant becomes a Debunker with a anti-rumor spreading rate β_5.
- Eq. (8): As time goes by, the rumor Debunkers lose their influence, and a Debunker becomes a Stifler_d with a anti-rumor fading rate β_6.
- Eq. (9): When a Spreader see the anti-rumor, the Spreader becomes a Debunker with a transformation rate β_7.
- Eq. (10): When a Stifler_r contacts a Debunker, the Stifler_r becomes a Debunker with a transformation rate β_8.
- Eq. (11): When a Stifler_r contacts a Debunker, the Stifler_r becomes a Stifler_d with a transformation rate β_9.

Following the classical rumor spreading model [30], our model starts with a homogeneous network. We assume that for a specific rumor, the population size is fixed, which is the whole population on the social media. Further more, the "infectious duration" and "infectious probability" of the population is homogeneous. These assumptions hold as we are studying the rumor spreading on social media in a macro level.

Let $S_1(t)$, $S_2(t)$, $S_3(t)$, $S_4(t)$, and $S_5(t)$ denote the percentages of the population that are Ignorants, Spreaders, Debunkers, Stiflers_r and Stiflers_d at time t, respectively. They satisfy the normalization condition: $S_1(t)+S_2(t)+S_3(t)+S_4(t)+S_5(t) = 1$. According to the Contact Eqs. (3)–(11), we take the derivatives of each group population with respect to time t, and the mean-field equations of the RSD system can be described as the following ODEs:

$$\begin{cases} \frac{dS_1(t)}{dt} = -\beta_1 kS_1(t)S_2(t) - \beta_3 kS_1(t)S_2(t) - \beta_4 kS_1(t)S_3(t) - \beta_5 kS_1(t)S_3(t), \\ \frac{dS_2(t)}{dt} = \beta_1 kS_1(t)S_2(t) - \beta_2 S_2(t) - \beta_7 kS_2(t)S_3(t), \\ \frac{dS_3(t)}{dt} = \beta_5 kS_1(t)S_3(t) - \beta_6 S_3(t) + \beta_7 kS_2(t)S_3(t) + \beta_8 kS_4(t)S_3(t), \\ \frac{dS_4(t)}{dt} = \beta_2 S_2(t) + \beta_3 kS_1(t)S_2(t) - \beta_8 kS_4(t)S_3(t) - \beta_9 kS_4(t)S_3(t), \\ \frac{dS_5(t)}{dt} = \beta_4 kS_1(t)S_3(t) + \beta_6 S_3(t) + \beta_9 kS_4(t)S_3(t), \end{cases} \quad (12)$$

where k denotes the average degree of the network. In Section 4, we use Runge–Kutta methods [36] to solve the differential equation system (12), and analyze the impacts of different modeling parameters (e.g., transition rates) on the rumor spreading–debunking process.

4. Data, case analyses, and results

4.1. Data

Hurricane Harvey of 2017 is tied with Hurricane Katrina of 2005 as the costliest tropical cyclone on record, inflicting $125 billion in damage, primarily from catastrophic rainfall-triggered flooding in the Houston metropolitan area and Southeast Texas. Harvey caused at least 107 confirmed deaths, displaced more than 30,000 people, and prompted more than 17,000 rescues [37].

We collected a real world data set during Hurricane Harvey in August, 2017 from Twitter.com. We focused on a popular rumor that is very important for the safety of the victims. After the hurricane happened, a rumor saying "Immigration status has to be checked before you are allowed to enter a shelter" started circulating through the social media. The rumor emerged because pictures were taken showing that U.S. Customs and Border Protection (CBP) Officers appeared in one of the shelter of Houston. This is a false rumor, because CBP officers were not checking the identity documents of the victims. However, if a victim who had illegal identity or did not bring the identity documents with himself saw this rumor, he might be afraid of getting caught and not choose to seek the shelter for protection. It could be dangerous for victims not coming to the shelters, and this rumor could cause severe loss of life and property. In this case, the official debunking started the debunking intervention timely, about one day after the rumor started getting popular.

To cover all related tweets, we collect all tweets that contain the keywords "shelter" and "immigration" since August, 2017. It is 2035 related Tweets collected during this event. Two independent coders manually coded all Tweets with label "Spreader", "Debunker", and "others". The "others" refers to some commentary Tweets or Tweets that are trying to seek confirmation about this rumor. Besides the time stamp and content, we collect the re-tweet counts of each original tweet.

The main variables in the model are the percentages of Spreaders and Debunkers that exist in each time step, which are $S_2(t)$ and $S_3(t)$ in the RSD system (12). These variables are measured as the sum of all individuals who post or re-tweet an original rumor/anti-rumor within each time interval. One hour is set as the unit of time in this paper.

As the exact time stamp of each re-tweet was not observable, we assume that after an original tweet was posted, the re-tweet delayed time probability distribution follows a beta distribution with shape parameters $\alpha = 2$, $\beta = 5$. The shape of the re-tweet time distribution is consistent with the beta distribution with $\alpha = 2$, $\beta = 5$ [38]. The curve first increases drastically and then decreases smoothly. We also assume that the re-tweeting behavior of the focal rumor will cease in one day because of information overload. For example, if the number of re-tweet of an original tweet is 100, we assume that starting from the posted time point, the re-tweet rate of this tweet increases drastically in the first few hours, and then decreases smoothly. We also assume that the tweet is re-posted 100 times within 24 h after the tweet was posted. Based on the assumptions discussed above, we measure the percentages of Spreaders ($S_2(t)$) and Debunkers ($S_3(t)$) that exist in each time step. Therefore, our data has provided the variation trend of the Spreaders and Debunkers. We will use this variation trend of Spreaders and Debunkers to derive the values of other parameters in the model, simulate the whole spreading process of the rumor case, and further estimate the scope of influence and more effective debunking strategies of the rumor.

4.2. A real world rumor case study: Emerging stage

Although our proposed model described by ODE system (1) applies to all rumor cases, the parameters in the model, such as the rumor spreading rate (β_1) and rumor fading rate (β_2), are distinct for different rumor cases. Our real life data is used to derive the case-specific parameters of this model. In this section, we will derive the case-specific parameters of the rumor emerging stage.

We design a RSD work-flow to determine the case-specific parameters (see Fig. 5). Firstly, a set of initial parameters are input into the system (1). By solving the ODE system using the initial parameters, we obtain the simulated variation

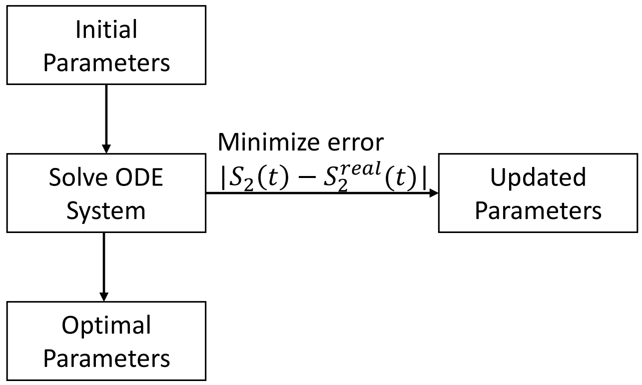


Fig. 5. Numerical implementation work-flow.

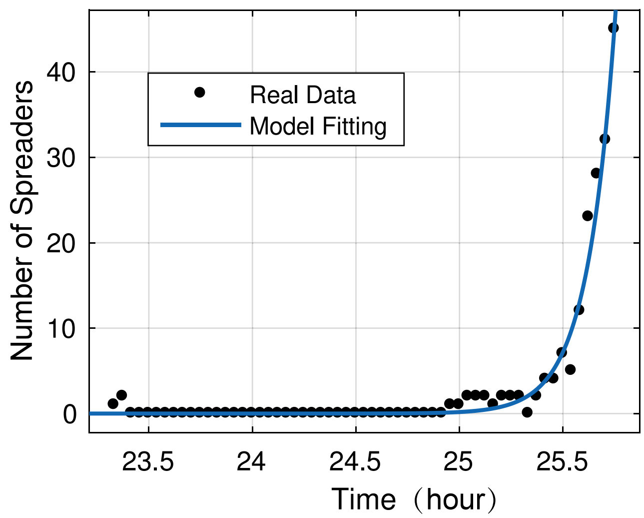


Fig. 6. Curve fitting of the real data.

trend of $S_2(t)$. Secondly, we calculate the error term $|S_2(t) - S_2^{real}(t)|$. Finally, by iteratively updating the parameters, we minimize the error term, and obtain the optimal set of parameters that best fits this rumor case.

The emerging stages is defined as the time period before the first official debunking information was released, which is August 25 at 7pm. The first rumor appeared as early as August 23 at 8 am. However, it has only one re-tweet and the rumor did not emerge again until August 25. Therefore, we use the analysis below to mathematically identify the starting time of this rumor case.

We assume that at the very beginning, only one spreader started this rumor. Hence, when $t = 0$, $S_1(0) = \frac{N-1}{N}$, $S_2(0) = \frac{1}{N}$, $S_3(0) = 0$, where N is the number of all active Twitter account in US. At the early stage of the rumor spreading, the number of Spreaders and the total number of Stiflers are small relative to the entire population N. Thus, the percentage of Ignorants is approximately invariant, i.e., $S_1 \approx 1$. By taking the integral of the second equation in (1), $S_2(t) \approx \frac{1}{N}e^{(\beta_1 k - \beta_2)(t+t_0)}$, and t_0 hours before August 23 at 8 am is the real starting time of this rumor case. The number of Spreaders is approximately $e^{(\beta_1 k - \beta_2)(t+t_0)}$. Note that when $\beta_1 k < \beta_2$, namely $\frac{\beta_1 k}{\beta_2} < 1$, the number of Spreaders is less than 1, and then the rumor ends; when $\beta_1 k > \beta_2$, namely $\frac{\beta_1 k}{\beta_2} > 1$, the number of Spreaders is gradually blooming, and then the rumor is uncontrollable. The result is consistent with Proposition 2. We used Matlab as a tool to apply the curve fitting of the real data to minimize $|S_2(t) - S_2^{real}(t)|$ (See Fig. 6). The dot is the data point of number of Spreaders in each time point, and the line is the fitted function curve, which is $e^{(\beta_1 k - \beta_2)(t+t_0)} \approx e^{0.3076(t-46)}$. Hence, we obtain the estimated parameter $\beta_1 k - \beta_2 = 0.3076$, and $t_0 = -46$. Hence, 46 h after August 23 at 8 am is the real starting time of this rumor case, which is August 25 at 6 am. This results indicated that the rumor had been circulating for 13 h before the first official debunking information released. The result shows a good match with the real data with a goodness of fit R-square = 0.98.

For the Immigration Rumor, we define the entire population as non-robot daily active Twitter users in US ($N = 27,744,000$). N is derived from the

$$N = a * p * r,$$

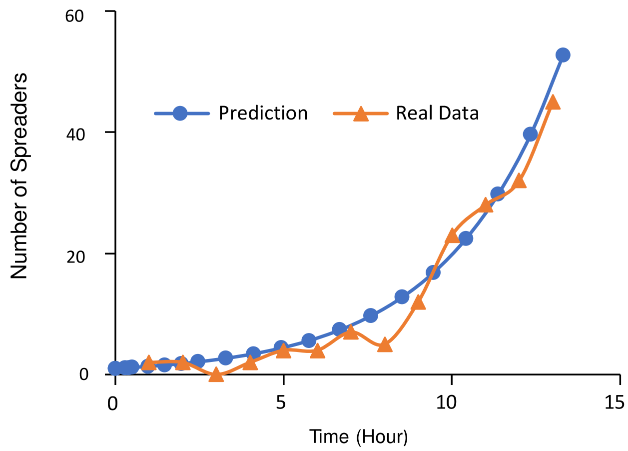


Fig. 7. Comparison of the number of spreaders between simulated and real data.

where a is the number of total active monthly users in US ($a = 68,000,000$),[2] p is the percentage of daily users ($p = 48\%$),[3] and r is the percentage of non-robot ($r = 85\%$) users on Twitter.[4] Because of the information overload in current network, we assume that the influence of one tweet would fade after 24 h. Hence, the fading rate

$$\beta_2 = \frac{1}{\Delta t} = \frac{1}{24} = 0.04.$$

As $\beta_1 k - \beta_2 = 0.3076$, the spreading rate $\beta_1 k = 0.34$. Finally, the impression rate β_3 is derived from

$$\beta_3 k = \frac{\text{total impressions in one hour}}{N} = \frac{f * p * r}{N * \Delta t} = \frac{1\,511\,626 \times 48\% \times 85\%}{27\,744\,000 \times 24} = 0.0009,$$

where $f = 1,511,626$ is the sum of the total followers of all rumor Spreaders in our data set. These people are the whole population that are possible to see the rumor.

We then substitute the estimated case-specific parameter to the ODE system (1), $\beta_1 k = 0.34$, $\beta_2 = 0.04$, $\beta_3 k = 0.0009$, $N = 27,744,000$, $S_1(0) = \frac{N-1}{N}$, $S_2(0) = \frac{1}{N}$, $S_3(0) = 0$. We use the Runge–Kutta method to derive the solution of the system. Comparing to the number of Spreaders in our data set, we can see that simulated trend of number of Spreaders matches well with the real trend (See Fig. 7).

Validated by the real life data, our model could be used to predict the future trend of this rumor. Using the estimated parameter sets, we could simulate and calculate $S_1(t)$, $S_2(t)$, and $S_3(t)$ at any point of time t. In this section, we conduct two simulations of the focal rumor. Firstly, we estimate the rumor spreading expansion trend if no debunking intervention is taken. As shown in Fig. 8(a), without debunking intervention, the rumor will keep spreading until the whole population were affected by the rumor.

Secondly, we estimate the total number of people that are influenced by the rumor. As decision makers could not observe the number of Stiflers (i.e., the amount of people who already know about the rumor), the model could be used to estimate the current scope of the rumor (See Fig. 8(b)). In our rumor case, by the time when the first anti-rumor appears, the scope of the rumor is approximately 340,000 Twitter accounts.

Therefore, using the stage 1 of our model, decision makers could draw two conclusions when they notice the rumor. First, if they do not take any debunking action, the rumor will keep spreading until all Twitter users are aware of the rumor. Second, by the time when they find out the rumor, it is estimated that 340,000 people on the Twitter are aware of the rumor. Measures could be taken accordingly to intervene this rumor case.

4.3. A real world rumor case study: Debunking stage

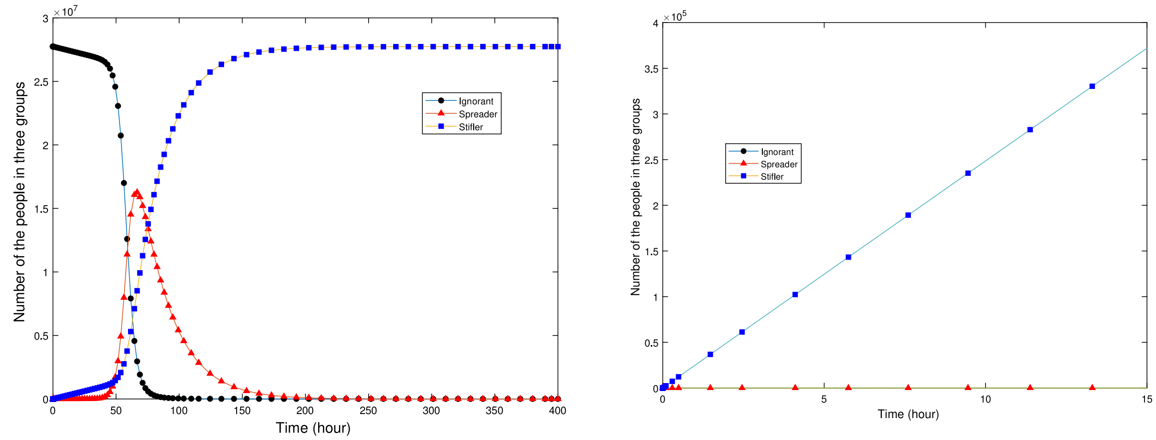
At the Debunking Stage, the rumor will keep spreading through the social media. At the same time, the anti-rumor starts circulating. The spreading of anti-rumor will interplay with the rumor spreading process, transferring Ignorants, Spreaders, and Stiflers_r into Debunkers and Stiflers_d. Fig. 9 shows how the number of Spreaders and Debunkers vary over time during the Hurricane Harvey immigration rumor.

Similarly to Section 4.2, we first estimate the rumor-specific parameters β_i, $i = 1, 2, \ldots, 9$, of the immigration rumor. Secondly, we implemented the derived set of parameters to solve the differential equations. Finally, we compared the simulated solution of the ODE system (12) with the real data to validate our model.

[2] https://www.statista.com/statistics/232818/active-us-twitter-user-growth/

[3] https://www.omnicoreagency.com/twitter-statistics/

[4] https://www.cnbc.com/2017/03/10/nearly-48-million-twitter-accounts-could-be-bots-says-study.html



(a) Rumor Spreading Lifecyle without Debunkers (b) The number of Stiflers at Stage 1

Fig. 8. Simulations of the immigration rumor at Stage 1.

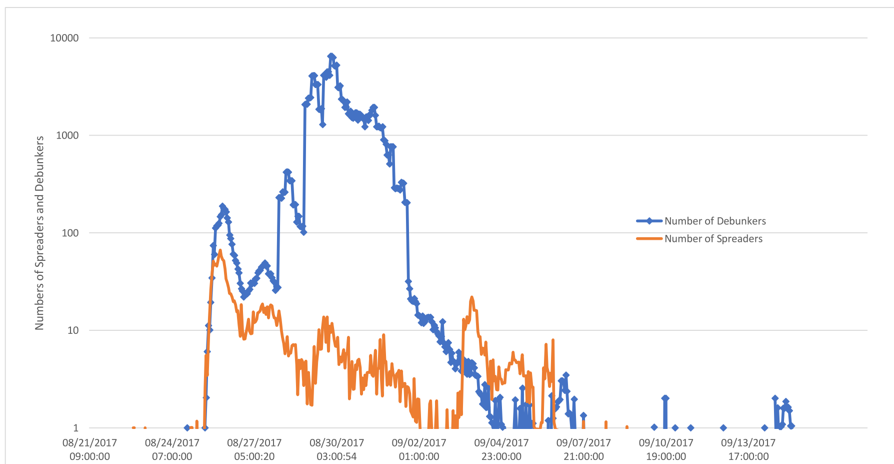


Fig. 9. Timeline of number of Spreaders and Debunkers in the Hurricane Harvey immigration rumor Aug. 21, 2017–Sep. 13, 2017.

We denote the starting time of the Debunking stage as t_0. The number of existing Spreaders i and Stiflers_r j at time t_0 has been derived from the stage 1. Since the first official Debunker appears at time t_0, we have $S_1(t_0) = \frac{N-i-j-1}{N}$, $S_2(t_0) = \frac{i}{N}$, $S_3(t_0) = \frac{1}{N}$, $S_4(t_0) = \frac{j}{N}$, and $S_5(t_0) = 0$.

By the results of Stage 1, at time t_0, the number of Spreaders $i = 50$, and the number of Stiflers $j = 340,000$. Then we use the RSD work-flow to determine the optimal sets of the parameters. For the initial parameters, we assume that the popularity of the anti-rumor spreading show the same pattern as that of rumor spreading. Therefore, the spreading rate, fading rate, and the impression rate of the rumor and anti-rumor spreading are comparative, i.e., $\beta_5 k = \beta_1 k = \beta_7 k = \beta_8 k = 0.34$, $\beta_6 = \beta_2 = 0.04$, $\beta_4 k = \beta_3 k = \beta_9 k = 0.0009$. Method of bisection are used to get updated β_i that minimizing $|S_2(t) - S_2^{real}(t)| + |S_3(t) - S_3^{real}(t)|$. The optimal sets of β_i are $\beta_1 k = 0.02$; $\beta_2 = 0.02$, $\beta_3 k = 0.0009$, $\beta_4 k = 0.0009$, $\beta_5 k = 0.15$, $\beta_6 = 0.04$, $\beta_7 k = 0.2$, $\beta_8 k = 0.4$, $\beta_9 k = 0.0009$. The comparison of simulated curve with the derived parameters matches well with our real data (see Fig. 10).

The model of the second stage will help us to understand the Debunking process more clearly. Besides the variation trend of the Spreaders and Debunkers, we could further estimate the variation trend of the Ignorants, Stiflers_r, and Stiflers_d. As shown in Fig. 11, most of the active Twitter users during that period of time would be aware of the correct news, namely that the immigration status will not be checked in the shelter. However, the number of Stiflers_r are 33,362, which means the number of people who know about the rumor but are still unaware of the correct news is around 33,362. In this case, the rumor debunking is timely and efficient. The proportion of uncleared Twitter users are small relative to the number of all impacted people.

In the next section, we will use sensitivity analysis as a tool to explore how different parameters and Debunking measures will impact the debunking effectiveness in this case.

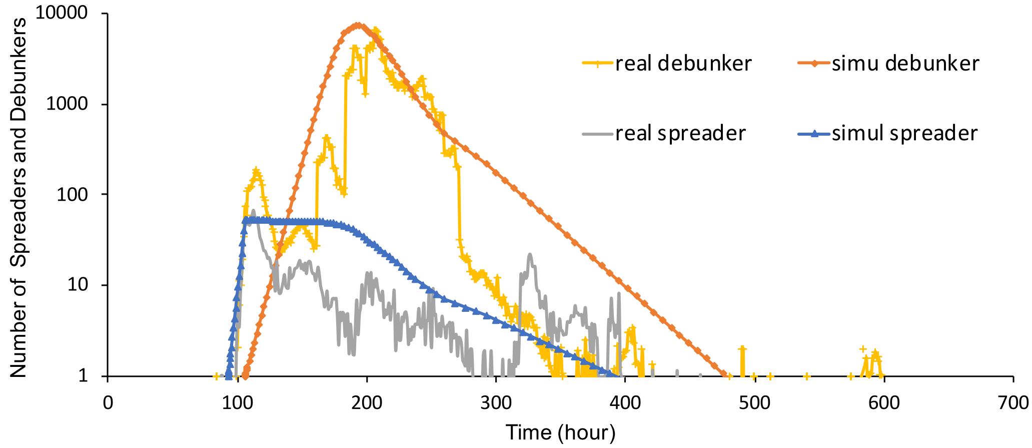


Fig. 10. Comparison of simulated vs. real Spreaders and Debunkers.

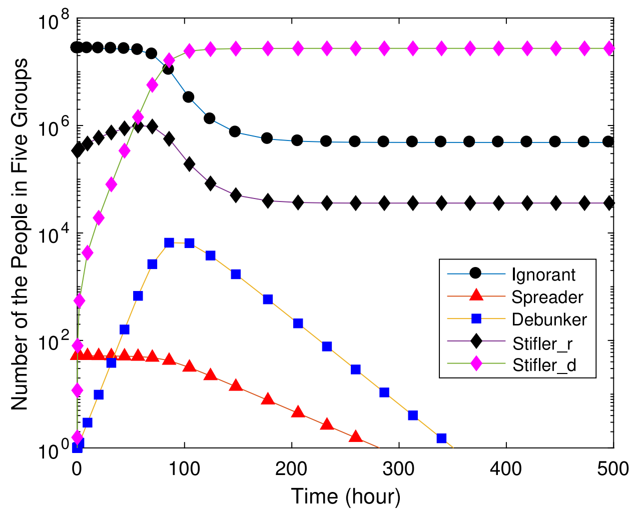


Fig. 11. Estimated variation trend of the Five State of Twitter users.

4.4. Sensitivity analysis

Sensitivity analysis is applied to identify the criticality of the modeling parameters and the effectiveness of debunking strategies in the rumor debunking process. This could help us to propose better rumor debunking strategies and diminish the impact of misinformation in the further cases.

Firstly, we compare the estimated final scope of the rumor under the situations with and without Debunkers in this case, and how the rumor final scope vary over the rumor popularity (β_1). As shown in Fig. 12, the final scope of the rumor is similar under the conditions with and without Debunkers when the rumor is not very popular ($\beta_1 < 0.05$). As the popularity of the rumor case increases, the difference of the uncleared Twitter users between the conditions with and without the Debunker first increases exponentially and then becomes stabilized. This difference becomes over 10^7 when $\beta_1 > 0.06$, which means the debunking intervene play a crucial role for the popular rumors.

Secondly, we compare how the number of Spreaders and the final scope of the rumor vary over different numbers of initial Debunkers at time t_0. As shown in Fig. 13(a), as the number of the Debunkers increases, the number of Spreaders decreases faster. The number of Stiflers_r, which can measures the rumor final scope, decreases from 32,500 to 29,000 when 18 initial Debunkers were added (see Fig. 13(b)). However, adding 180 more Debunkers only decreases the final scope by approximately 5000, which may not be worth for the investigated human resources. We define the equilibrium point of rumor spreading as the time point when the number of Spreaders is less than 1. At this time-point, the rumor stops spreading. We compare how the equilibrium points vary as different numbers of Debunkers appearing at time t_0. From Fig. 13(a) and (b), we observe that the more initial Debunkers are at t_0, the earlier this rumor ends. Hence, we are interested to study the end time of the rumor case. This is also a very important variables, as government officials or companies all want to get rid of the influence of the misinformation as soon as possible. The end time is defined as the point of time that the system reaches equilibrium point, and the number of spreaders become less than 1. Fig. 13(c)

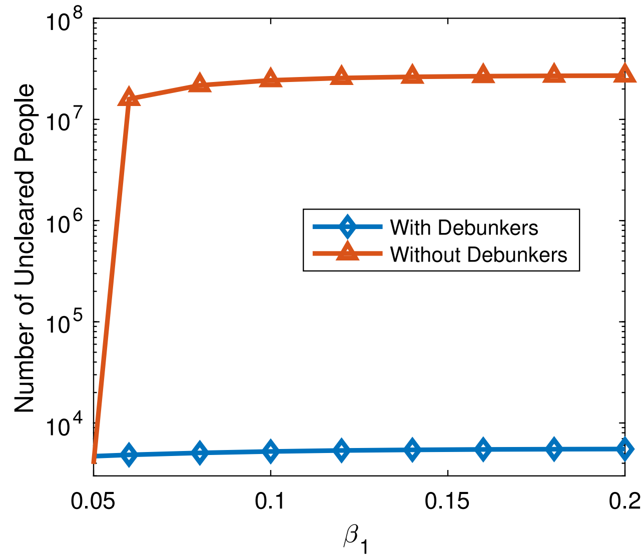


Fig. 12. Final scale of rumor with & without Debunkers.

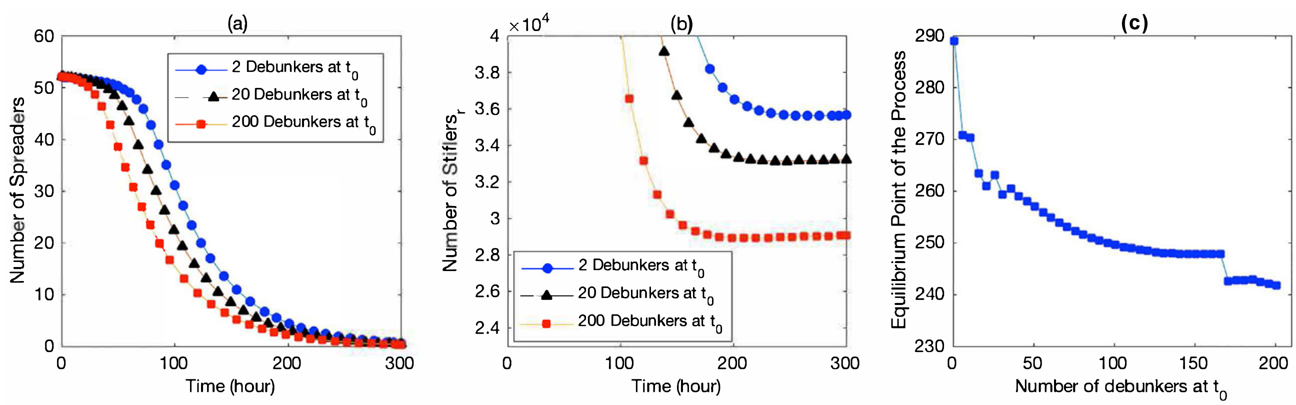


Fig. 13. Sensitivity analysis when varying number of Debunkers at time $\mathbf{t}_0$.

shows that the equilibrium point is decreasing, with diminishing marginal returns while adding more Debunkers. This result is consistent with the results in Fig. 13(a) and (b). Government officials or companies could determine their invested debunking human resources to achieve their desired debunking results according to this result. Clearly, the optimal number of Debunkers in Fig. 13(c) is 20, and the sub-optimal number is 170.

Thirdly, we investigate how the end time varies over the number of initial Spreaders. From Fig. 14, the duration of the rumor increases concavely as the number of initial spreader increases.

Finally, we discuss the impact of the Debunking starting time on the rumor spreading process. We simulate four different intervention time: $t_0 = 6$, $t_0 = 12$, $t_0 = 24$, $t_0 = 36$, where t_0 is the number of hours after the rumor starts circulating. As shown in Fig. 15, if the rumor spreading process is debunked within 6 h, the rumor will completely end within 180 h; if officials delayed debunking intervention by only 6 h, the end time will be delayed by 100 h when debunking time $t_0 = 12$. Therefore, it is important for officials to take action as soon as possible to reduce the loss and influence of the rumor.

5. Discussion

In this paper, we use an epidemiological method to study the spreading–debunking process of a rumor. We divide the whole life-cycle of a rumor into two stages. The first stage is emerging stage, which refers to the time period between the rumor starting point and the official debunking point. In the second stage, with the release of official debunking information, people start to spread the anti-rumor in the social network. The spreading of rumors and anti-rumors is a dynamic and reciprocal process.

In the first stage, we study the rumor spreading pattern before Debunkers appear. Before the appearance of the first debunker and/or verification information, the system consists of Spreaders, Ignorants and Stiflers. We analytically study

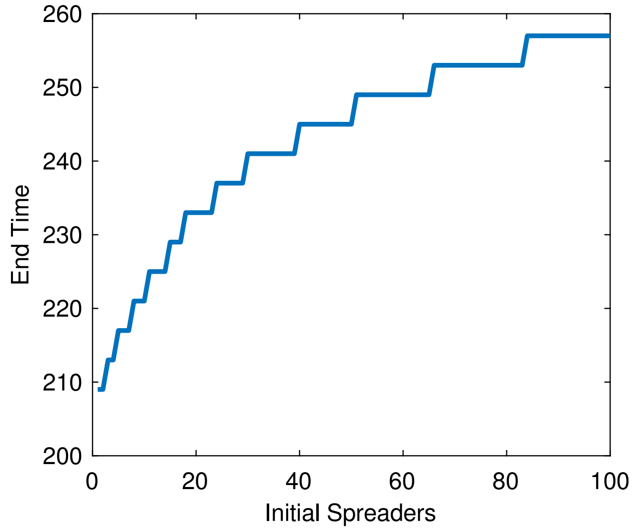


Fig. 14. Rumor duration varies over number of initial spreaders.

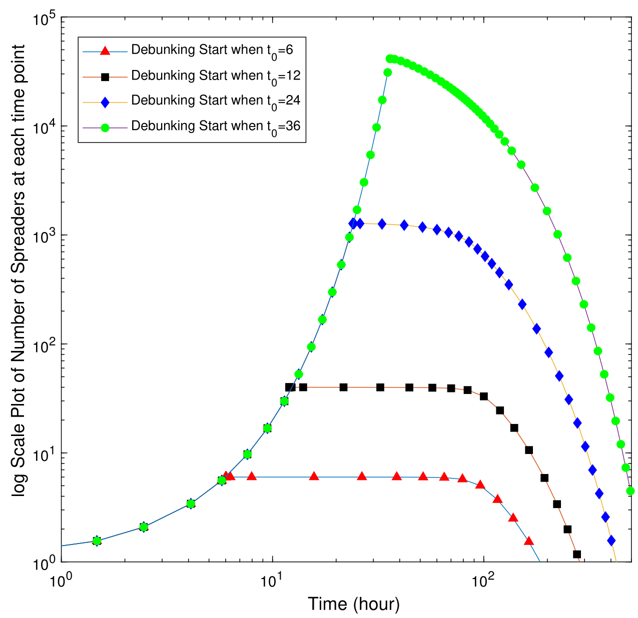


Fig. 15. The impact of Debunk starting time on the Rumor Spreading process.

the properties of the rumor, including its final scope and maximal scale, and how it will vary over the spreading rate and fading rate. In the real world case study, we validate our model with the real data. The spreading trend and current scope of the Rumor (how many people have already heard about the rumor) is estimated through the model.

For the second stage, we built a novel model to explore the interplay mechanism between rumor spreading and debunking processes. We solve the system and compared them with the real data. The real data matches well with our predicted solution. From these results, we could understand how many people are influenced by the rumor, and are still not aware of the real information.

Moreover, four sets of sensitivity analysis are applied. We conclude that the debunking intervention is not necessary when the rumor is not very popular with $\beta_1 < 0.05$. However, when $\beta_1 > 0.06$, the debunking intervention is critical and could make a significant influence on the final scope of the rumor. Secondly, we conclude that the end time decreases with diminishing marginal returns while adding more Debunkers. Government officials or companies could determine their invested debunking human resources to achieve their desired debunking results according to our result. Thirdly, the duration of the rumor increases concavely as the number of initial spreader increases. Finally, we discussed the impact of the Debunking starting time on the rumor spreading process. The result shows that it is important for officials to take action as soon as possible to reduce the loss and influence of the rumor.

CRediT authorship contribution statement

Meiling Jiang: Software, Methodology, Validation, Data curation, Writing - original draft. **Qingwu Gao:** Methodology, Formal analysis, Investigation, Writing - original draft, Writing - review & editing. **Jun Zhuang:** Conceptualization, Methodology, Resources, Visualization, Supervision, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix

Proof of Proposition 1. Dividing Eq. (1) by Eq. (3), we can get

$$\frac{dS_1(t)}{dS_3(t)} = -\frac{S_1(t)(\beta_1+\beta_3)k}{\beta_2+\beta_3 k S_1(t)},$$

which leads to

$$-dS_3(t) = \frac{\beta_2+\beta_3 k S_1(t)}{S_1(t)(\beta_1+\beta_3)k} dS_1(t),$$

or, equivalently,

$$-S_3(t) = \frac{1}{(\beta_2+\beta_3)k}(\beta_3 k S_1(t) + \beta_2 \ln S_1(t)).$$

Therefore, $S_1(t)$ and $S_3(t)$ satisfy the implicit solution equation

$$\beta_3 k S_1(t) + \beta_2 \ln S_1(t) + (\beta_1+\beta_3)kS_3(t) = c,$$

where c is a constant determined by the initial condition. We assume that there is only one spreader who starts the rumor at the initial stage of the rumor. $S_1(0) = \frac{N-1}{N} \approx 1$, $S_2(0) = \frac{1}{N} \approx 0$, $S_3(0) = 0$. Substitute the values to the implicit solution, we have $c = \beta_3 k$.

We define the end point of the rumor as the point when the number of spreaders is less than 1. Hence, $S_2(\infty) \approx 0$. As $S_1(t)+S_2(t)+S_3(t) = 1$, $S_1(\infty) = 1 - S_3(\infty)$.

Therefore, at the end point when $t \to \infty$,

$$\beta_3 k(1-S_3(\infty)) + \beta_2 \ln(1-S_3(\infty)) + (\beta_1+\beta_3)kS_3(\infty) = \beta_3 k,$$

which implies that

$$\ln(1-S_3(\infty)) = -\frac{\beta_1 k}{\beta_2} S_3(\infty),$$

namely,

$$S_3(\infty) = 1 - e^{-\frac{\beta_1 k}{\beta_2} S_3(\infty)}.$$

Let $R = S_3(\infty)$, Eq. (2) is satisfied.

Clearly, $R = 0$ is a trivial solution of Eq. (2). Now in the rest part, we prove the existence of the non-trivial solution R, $0 < R < 1$, under the condition that $\frac{\beta_1 k}{\beta_2} > 1$. Set $f(x) = x - 1 + e^{-\frac{\beta_1 k}{\beta_2} x}$, $0 < x < 1$. Then $f'(x) = 1 - \frac{\beta_1 k}{\beta_2} e^{-\frac{\beta_1 k}{\beta_2} x}$, and $f'(x_0) = 0$ for $x_0 = \frac{\beta_2}{\beta_1 k} \ln \frac{\beta_1 k}{\beta_2}$, where $x_0 \in (0,1)$ since $\frac{\beta_1 k}{\beta_2} > 1$. Moreover, $f(1) = e^{-\frac{\beta_1 k}{\beta_2}} > 0$, and $f(x_0) = \frac{\beta_2}{\beta_1 k}\left(\ln \frac{\beta_1 k}{\beta_2} - \frac{\beta_2}{\beta_1 k} + 1\right)$, where $f(x_0) < 0$ since $\ln x < x - 1$ for $x \in (0,1)$. By the Intermediate Value Theorem, equation $f(x) = 0$ has a non-trivial solution R such that $0 < x_0 < R < 1$, which shows that for $\frac{\beta_1 k}{\beta_2} > 1$ there exists the claimed non-trivial solution R, $0 < R < 1$. □

Proof of Proposition 2. Without loss of generality, we only need to prove that R increases as β_1 increases. Taking the partial derivative of R with respect β_1 on the both sides of Eq. (2), we have

$$\frac{\partial R}{\partial \beta_1} = e^{-\frac{\beta_1 k}{\beta_2} R}\left(\frac{k}{\beta_2} R + \frac{\beta_1 k}{\beta_2}\frac{\partial R}{\partial \beta_1}\right),$$

which yields that

$$\left(1 - \frac{\beta_1 k}{\beta_2} e^{-\frac{\beta_1 k}{\beta_2} R}\right)\frac{\partial R}{\partial \beta_1} = \frac{k}{\beta_2} R e^{-\frac{\beta_1 k}{\beta_2} R}.$$

Further, we have

$$\frac{\partial R}{\partial \beta_1} = \frac{\frac{k}{\beta_2} R e^{-\frac{\beta_1 k}{\beta_2} R}}{1 - \frac{\beta_1 k}{\beta_2} e^{-\frac{\beta_1 k}{\beta_2} R}}.$$

Let $u = \frac{\beta_1 k}{\beta_2}$. In the following, it suffices to show that $1 - ue^{-uR} > 0$. Note that R is the nontrivial solution of Eq. (2) and also is a function of β_1. Thus, point R (more formally, point $(R, 0)$) is the point on the horizontal axis corresponding to the intersection (not the origin) of line $y = x$ and curve $y = 1 - e^{-ux}$. Since $1 - e^{-uR} > 1 - ue^{-uR}$, curve $y = 1 - e^{-uR}$ is above curve $y = 1 - ue^{-uR}$; curve $y = 1 - ue^{-uR}$ cross the horizontal axis at $x = \frac{\ln u}{u}$. Then we can show that $1 - ue^{-uR} > 0$, if we prove that on the horizontal axis point $x = \frac{\ln u}{u}$ is on the left side of $x = R$, in other words, $1 - e^{-u(\ln u)/u} > \frac{\ln u}{u}$, i.e. $1 - \frac{1}{u} - \frac{\ln u}{u} > 0$. In fact, define a function $f(u) = 1 - \frac{1}{u} - \frac{\ln u}{u}$. Since $f'(u) = \frac{\ln u}{u^2} > 0$ and $f(1) = 0$, we obtain $f(u) > 0$ for any $u > 1$. □

Proof of Proposition 4. Dividing Eq. (2) by Eq. (1), we have

$$\frac{dS_2(t)}{dS_1(t)} = \frac{\beta_2}{(\beta_1 + \beta_3)kS_1(t)} - \frac{\beta_1}{\beta_1 + \beta_3}, \tag{13}$$

namely,

$$dS_2(t) = \left(\frac{\beta_2}{(\beta_1 + \beta_3)kS_1(t)} - \frac{\beta_1}{\beta_1 + \beta_3}\right) dS_1(t).$$

Then, we take the integral on both sides of the above equation to obtain that

$$S_2(t) = \frac{\beta_2}{(\beta_1 + \beta_3)k} \ln S_1(t) - \frac{\beta_1}{\beta_1 + \beta_3} S_1(t) + C, \tag{14}$$

where the constant C is determined by the initial conditions:

$$S_1(0) = \frac{N-1}{N} \approx 1, \quad S_2(0) = \frac{1}{N} \approx 0 \ \text{ and } \ S_3(0) = 0,$$

following from the fact that there is only one spreader who starts the rumor at the initial stage of the rumor. Hence, we substitute the initial conditions into (14) to get

$$C = \frac{\beta_1}{\beta_1 + \beta_3}.$$

From (14), we derive that $S_2(t)$ has the explicit expression with respect to $S_1(t)$ as

$$S_2(t) = \frac{\beta_2}{(\beta_1 + \beta_3)k} \ln S_1(t) - \frac{\beta_1}{\beta_1 + \beta_3} S_1(t) + \frac{\beta_1}{\beta_1 + \beta_3}. \tag{15}$$

From (13), letting $dS_2/dS_1 = 0$ implies that

$$S_1 = \frac{\beta_2}{\beta_1 k},$$

at which we get the maximum scale of the spreaders as:

$$S_{\max} = \frac{\beta_2}{(\beta_1 + \beta_3)k} \left(\ln \frac{\beta_2}{\beta_1 k} - 1\right) + \frac{\beta_1}{(\beta_1 + \beta_3)}. \quad \square$$

References

[1] S. Aral, D. Walker, Creating social contagion through viral product design: A randomized trial of peer influence in networks, Manage. Sci. 57 (9) (2011) 1623–1639.

[2] G. Shi, A. Proutiere, M. Johansson, J.S. Baras, K.H. Johansson, The evolution of beliefs over signed social networks, Oper. Res. 64 (3) (2016) 585–604.

[3] M.C. Moldoveanu, J.A. Baum, I think you think I think you're lying: The interactive epistemology of trust in social networks, Manage. Sci. 57 (2) (2011) 393–412.

[4] L. Howell, et al., Digital wildfires in a hyperconnected world, WEF Report 3 (2013) 15–94.

[5] S.-Y. Lee, L. Qiu, A. Whinston, Sentiment manipulation in online platforms: An analysis of movie tweets, Prod. Oper. Manage. 27 (3) (2018) 393–416.

[6] N. Bennett, 7 false rumors about Harvey, 2017, URL http://www.khou.com/news/local/fact-vs-fiction-hurrican-harvey-rumors/468792969.

[7] F. Alfonso, Corpus christi debunks a second Hurricane Harvey rumor, 2017, URL http://www.chron.com/news/houston-weather/hurricaneharvey/article/hurricane-harvey-rumor-fake-false-facebook-corpus-11982732.php.

[8] K. Hunt, B. Wang, J. Zhuang, et al., Misinformation debunking and cross-platform information sharing through Twitter during Hurricanes Harvey and Irma: A case study on shelters and ID checks, Nat. Hazards (2020) 1–23.

[9] M.O. Jackson, Networks and economic behavior, Annu. Rev. Econ. 1 (1) (2009) 489–511.
[10] L. Zhao, J. Wang, Y. Chen, Q. Wang, J. Cheng, H. Cui, SIHR rumor spreading model in social networks, Physica A 391 (7) (2012) 2444–2453.
[11] H. Rahmandad, J. Sterman, Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models, Manage. Sci. 54 (5) (2008) 998–1014.
[12] W.A. Peterson, N.P. Gist, Rumor and public opinion, Am. J. Sociol. 57 (2) (1951) 159–167.
[13] R.H. Knapp, A psychology of rumor, Public Opin. Q. 8 (1) (1944) 22–37.
[14] J. Harsin, The rumor bomb: A convergence theory of contemporary mediated American politics, South. Rev. Commun. Polit. Cult. 39 (2006) 84–110.
[15] P. Bordia, Rumor Interaction Patterns on Computer-Mediated Communication Networks (Ph.D. dissertation), Temple University, 1996.
[16] A. Ganesh, L. Massoulié, D. Towsley, The effect of network topology on the spread of epidemics, in: 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE, Vol. 2, INFOCOM 2005, IEEE, 2005, pp. 1455–1466.
[17] O. Oh, K.H. Kwon, H. Rao, An exploration of social media in extreme events: Rumor theory and Twitter during the Haiti Earthquake 2010, in: ICIS, Vol. 231, 2010.
[18] O. Oh, M. Agrawal, H.R. Rao, Community intelligence and social media services: A rumor theoretic analysis of tweets during social crises, MIS Q. 37 (2) (2013) 407–426.
[19] K. Lerman, T. Hogg, Using a model of social dynamics to predict popularity of news, in: Proceedings of the 19th International Conference on World Wide Web, ACM, 2010, pp. 621–630.
[20] Z. Zhang, Z. Zhang, An interplay model for rumour spreading and emergency development, Physica A 388 (19) (2009) 4159–4166.
[21] K. Zhu, L. Ying, Information source detection in the SIR model: A sample path based approach, in: Information Theory and Applications Workshop, 2013, pp. 1–9.
[22] J. Huang, X. Jin, Preventing rumor spreading on small-world networks, J. Syst. Sci. Complex. 24 (3) (2011) 449–456.
[23] J. Lee, M. Agrawal, H. Rao, Message diffusion through social network service: The case of rumor and non-rumor related tweets during Boston bombing 2013, Inf. Syst. Front. 17 (5) (2015) 997–1005, http://dx.doi.org/10.1007/s10796-015-9568-z.
[24] R.-Y. Tian, Y.-J. Liu, Isolation, insertion, and reconstruction: three strategies to intervene in rumor spread based on supernetwork model, Decis. Support Syst. 67 (2014) 121–130.
[25] D. Shah, T. Zaman, Finding rumor sources on random trees, Oper. Res. 64 (3) (2016) 736–755.
[26] J. Zhou, Y.-J. Chen, Targeted information release in social networks, Oper. Res. 64 (3) (2016) 721–735.
[27] A. Cintron-Arias, Modeling and Parameter Estimation of Contact Processes (Ph.D. dissertation), Cornell University, 2006.
[28] C. Lefévre, Optimal control of a birth and death epidemic process, Oper. Res. 29 (5) (1981) 971–982.
[29] Wikipedia, Compartmental models in epidemiology–Wikipedia, The Free Encyclopedia, 2019, URL https://en.wikipedia.org/w/index.php?title=Compartmental_models_in_epidemiology&oldid=788424012. Online; (Accessed 11 March 2019).
[30] D. Kendall, et al., Stochastic rumors, Inst. Math. Appl. 1 (1965) 42–55.
[31] D.P.T. Maki, et al., Mathematical Models and Applications: With Emphasis on the Social Life, and Management Sciences, Technical report, 1973.
[32] C.F. Coletti, P.M. Rodríguez, R.B. Schinazi, A spatial stochastic model for rumor transmission, J. Stat. Phys. 147 (2) (2012) 375–381.
[33] D.H. Zanette, Dynamics of rumor propagation on small-world networks, Phys. Rev. E 65 (4) (2002) 041908.
[34] J. Wang, L. Zhao, R. Huang, 2SI2R rumor spreading model in homogeneous networks, Physica A 413 (2014) 153–161.
[35] B. Wang, J. Zhuang, Crisis information distribution on Twitter: a content analysis of tweets during Hurricane Sandy, Nat. Hazards (2017) 1–21.
[36] J.C. Butcher, The Numerical Analysis of Ordinary Differential Equations: Runge-Kutta and General Linear Methods, Wiley-Interscience, 1987.
[37] Wikipedia, Hurricane Harvey–Wikipedia, The Free Encyclopedia, 2019, URL https://en.wikipedia.org/wiki/Hurricane_Harvey. Online; (Accessed 02 February 2019).
[38] E. Spiro, C. Irvine, C. DuBois, C. Butts, Waiting for a retweet: modeling waiting times in information propagation, in: 2012 NIPS Workshop of Social Networks and Social Media Conference, Vol. 12, 2012, Accessed, http://Snap.Stanford.Edu/Social2012/Papers/Spiro-Dubois-Butts.Pdf.